DOCUMENT RESUME

ED 130 618 95 IR 004 149

AUTHOR Brown, O. Robert, Jr.
TITLE A Comparison of Test Scores of Students Using Programed Instructional Materials with Those of Students Not Using Programed Instructional Materials. Technical Report No.3.
INSTITUTION Illinois Univ., Urbana. Committee on School Mathematics.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Educational Media Branch.
REPORT NO P-711151.01
PUB DATE Jul 62
NOTE 25p.; Archival document

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Achievement; *Comparative Analysis; *Conventional Instruction; Mathematics Instruction; *Programed Instruction; Programed Materials; Secondary Education; *Secondary School Mathematics

ABSTRACT

During the first semester of the 1961-62 school year, students from eleven mathematics classes at seven high schools participated in a test of programed math materials. A control group composed of five classes was taught from a regular text by their math teachers; the experimental group contained six classes which were taught the same content via a combination of programed materials and classroom teacher. Achievement scores on seven unit tests were compared. At the beginning of the year all students took a test of general ability and a math Sequential Test of Educational Progress. On the ability tests the experimental group scored significantly higher than the control group. This difference was maintained for each unit test. No student studying from the programed materials was prevented from doing as well as he might have under more typical instructional procedures. Examination of cumulative records showed that the programed materials used were written for higher ability students. Item analysis showed that treatment had little effect on the ability to respond to recognition multiple-choice items, but had a significant effect on the ability to respond correctly to transfer items, in favor of the experimental group. (KB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

FD130618

UNIVERSITY OF ILLINOIS
Urbana, Illinois

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

COMPARATIVE STUDIES OF PRINCIPLES FOR PROGRAMMING MATHEMATICS IN AUTOMATED INSTRUCTION

Co-Investigators:

Lawrence M. Stolurow
Professor, Department of Psychology
Training Research Laboratory

Max Beberman
Professor, College of Education
University of Illinois Committee
on School Mathematics (UICSM)

Project Sponsor:

Educational Media Branch
U. S. Office of Education
Title VII

Project No. 711151.01

IR 004 149

U. S. Office of Education
Title VII

COMPARATIVE STUDIES OF PRINCIPLES FOR PROGRAMMING MATHEMATICS IN AUTOMATED INSTRUCTION

Technical Report No. 3

A Comparison of Test Scores of Students Using Programed Instructional Materials with those of Students Not Using Programed Instructional Materials

O. Robert Brown, Jr.

July, 1962

Purpose

This study summarizes some of the data gathered during the first semester of school year 1961-62. It is based on work from students participating in a trial-run of materials prepared by the UICSM Programed Instruction Project, and used for the first time. The students were from eleven classes at seven schools. The control group was composed of five classes which were taught from the regular text exclusively by their own mathematics teacher; the experimental group contained six classes which were taught the same content but via a combination of programed materials and classroom teacher. The performance of these groups on achievement tests is examined here.

Method

Materials

Programed Instructional Booklets. The programed instructional materials were prepared by UICSM project staff members. Each cooperating teacher in each group had demonstrated above average competence in previous teaching of the UICSM materials. The programed instructional materials used took the form of spiral-bound programed booklets, each booklet being called a 'Part'. The regular text was UICSM High School Mathematics, Unit 1.[1]

---

[1] Available through the University of Illinois Press, Champaign, Illinois.

Table 1

Experimental (E*) and Control (C) Test Means and SDs

Test	Means M E*	Means M C	Standard Deviations s E*	Standard Deviations s C
General Ability				
TOGA (IQ)	128.48	124.02	18.00	15.19
Math Aptitude				
STEP	37.28	34.62	6.13	6.69
Specific Achievement				
1	25.68	24.57	4.32	3.91
2	26.12	24.22	3.61	3.41
3	25.29	23.86[a]	4.56	4.64
4-Pre	23.68	23.64[b]	5.28	4.97
4-Post	29.36	28.11	2.73	2.38
5	24.33	22.01	4.91	4.19
6	25.67	24.25	4.21	4.56
7	25.74	24.38	3.51	3.52

[a]Control class S failed to take this test

[b]Control class T failed to take this test

It is not the intent of this paper to consider achievement test results in relation to the specific content covered. However, it may be of interest to the reader to know the topics which were covered by the programed materials and their related tests. Appendix D gives this information. Let us note again that students using the programed materials participated in regular classroom sessions for all topics of Unit 1 which were not programed. Consequently, they can be described as the group which was taught by the programed materials, a teacher, and conventional UICSM text; whereas, the other group, the control sample, was taught only by a teacher and conventional UICSM text.

Achievement Tests. A Part-test is simply a test of achievement over the content treated in a given Part of the programed material and in a related set of pages in the regular UICSM text. For example, Test 1 covers Part 1 of the programed materials and pages 1 through 6 of UICSM Unit 1.

Seven tests were designed, covering each of the first seven Parts; only one of these tests was given both as a pretest and as a post-test. On each test the student could earn a maximum score of 31 correct responses. Tests were administered as soon as possible after a student in the experimental group had completed a Part or as soon as a control class had completed the same content to its teacher's satisfaction. Control class teachers did not see the programed Parts; they had page references for Unit I to specify the content covered by the achievement tests.

Table 2

F ratios, t values, and Associated Probability Levels

Test	Groups Compared: E* and C		E* and C	
	F	p(F)	t	p(t)
General Ability				
TOGA (IQ)	1.40	.02<p<.1	2.44	<0.01
Math Aptitude				
STEP	1.09	---	3.68	<0.01
Specific Achievement				
1	1.22	---	2.42	<0.01
2	1.12	---	4.83	<0.01
3	1.04	---	2.66	<0.01
4-pre	1.13	---	0.07	---
4-post	1.12	---	3.98	<0.01
5	1.37	.02<p<.1	4.43	<0.01
6	1.17	---	3.00	<0.01
7	1.01	---	3.49	<0.01

Ability and Aptitude Tests. In addition to the seven Part-tests, each student took two other tests at the outset of the school year, a Test of General Ability[2] and a mathematics Sequential Test of Educational Progress,[3] referred to as TOGA and STEP, respectively.

Students

The control sample (C) was comprised of 183 students. The original experimental sample (E) was comprised of 193 students; one experimental class of 46 students was deleted from the data summaries presented here, since it was found to be so much lower in general ability level that there were no comparable control students.[4] Thus, the experimental sample (E*) as used included 147 of the original 193 students.

Results

Table 1 presents descriptive statistics computed for the two groups.

Table 2 presents t and F ratios based upon the measures described in Table 1, as well as the probability level associated with each ratio.

Some results from Table 2 should be specifically noted. In columns (1) and (3) both F and t ratios are computed for the comparisons of E* and C. On the ability tests, the E* Group was significantly better than the C Group with respect to both general IQ level and specific mathematical ability. This difference is maintained for each Part-test, except in the one pretest case, for which no significant difference is noticed. It appears that no student studying from the programed materials has been prevented from doing as well as he might have under more typical instructional procedures.

---

[2]Science Research Associates, Inc., Chicago, Illinois

[3]Educational Testing Service, Cooperative Test Division, Princeton, N. J.

[4]See Appendix A and Appendix B.

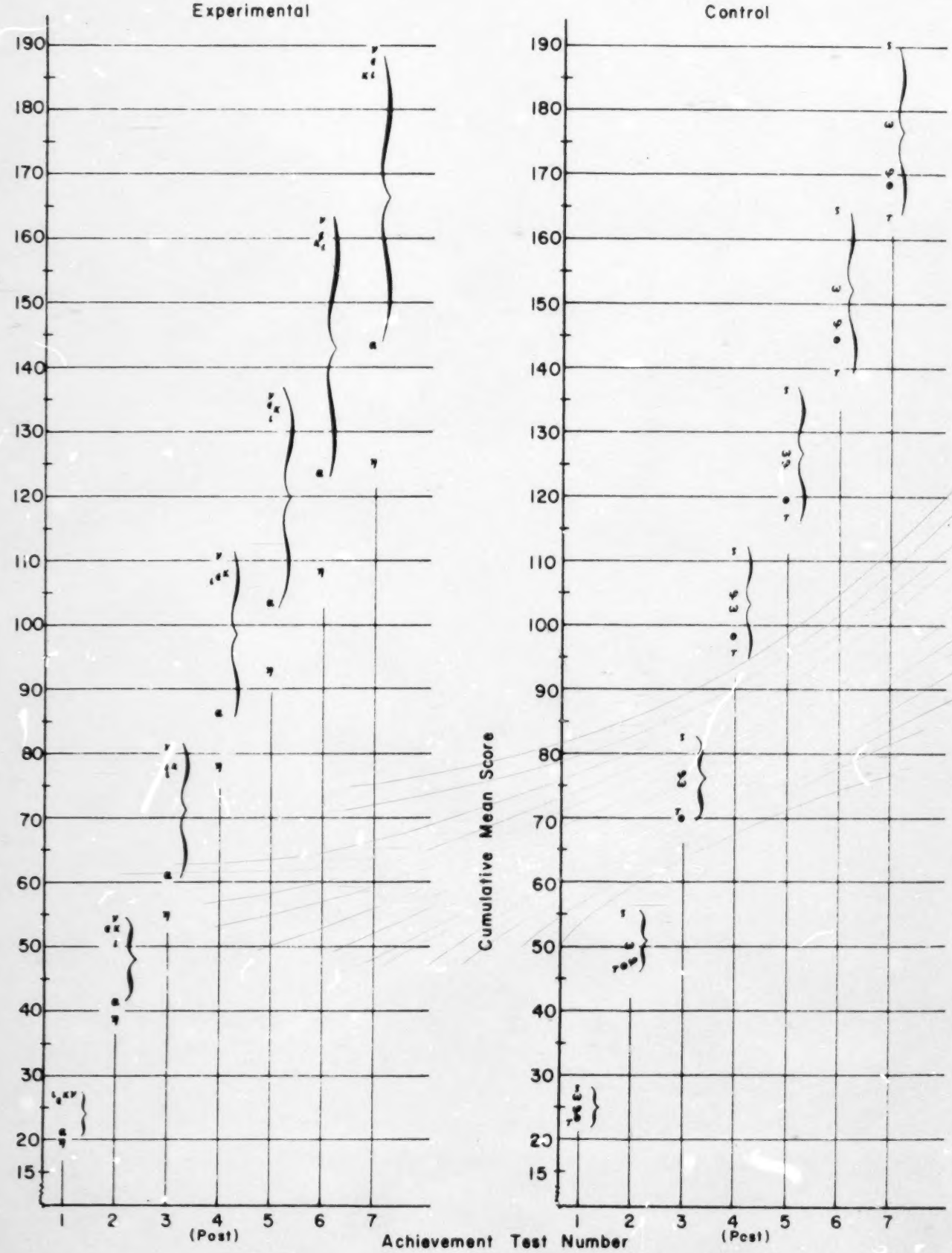


Fig. 1 Cumulative mean scores on achievement tests.

Another way to consider the effect of programed instruction is to look at a cumulative record. Figure 1 presents such a record by means of graphs treating E and C classes on independent axes, but using the same scale so that comparisons can be made. The graphs show the cumulative mean score after each test was administered. Each class is denoted on the graph by a lower case Greek letter.

Class η is the previously mentioned and eliminated lowest ability group. Class α also shows considerable deviation, probably due to its small number of students (N = 19) yet very wide ability range (IQ: 79-133, STEP: 13%-ile - 91%-ile). To show this class as an isolated unit is misleading for general observations.

A conclusion that may be drawn from the deviation of the η and α classes is that the programed materials that were used were written for higher ability students. That is, the authors, intentionally or not, directed their efforts at a segment of the school population higher in basic ability than the segment to which these two classes belong. [Appendix A gives the relevant ability data.]

It is interesting to note that although classes ϵ and ι differed significantly on their STEP mean [$t = -3.47$; $p < 0.01$], as did classes ι and κ [$t = 2.65$; $p < 0.01$], their cumulative mean scores are roughly the same. This condition does not occur among the control classes, in that the cumulative mean scores tend toward greater disparity as the number of tests increases.

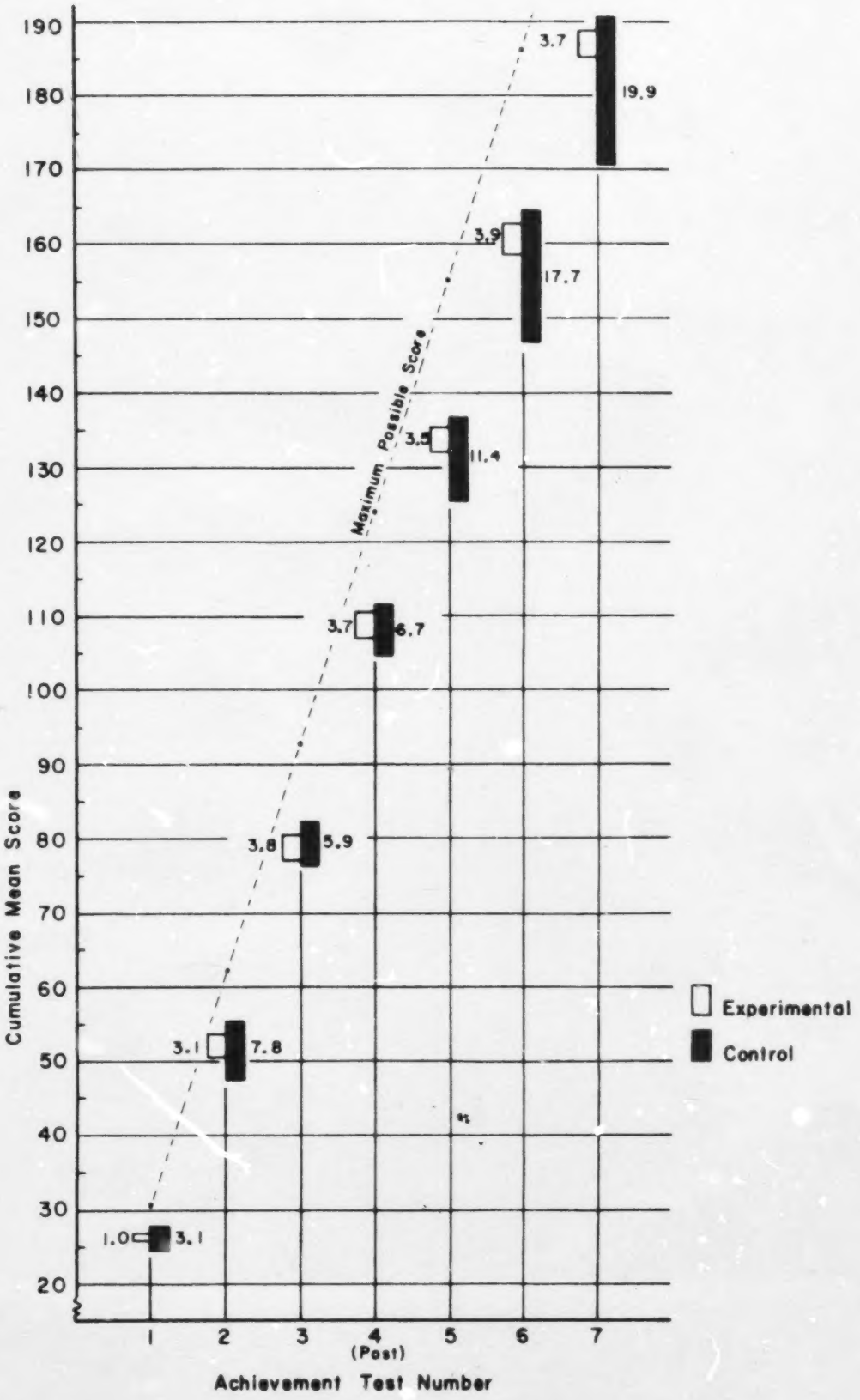


Fig. 2 Cumulative means.

Figure 2 shows the range of cumulative means of the four clustered E* classes together with the range of the same measure for the comparable C classes. If one were to project the results in Figure 2 over a greater number of Parts and related Part-tests, it seems clear that the programed material would be more successful in maintaining more homogeneous performance levels. It might be inferred that the original disparity under conventional instructional procedures would probably continue to be evidenced since it seems to be increasing.

Item Analyses

The second section of this report deals with findings obtained through the item analyses of achievement tests. Table 3 shows estimates of the internal reliability coefficient for each test, based on item difficulty levels. The method of computation assumes that each test is a power test [75% of the subjects answered all items] and that item difficulties are reasonably uniform.[5]

The reliability levels are reasonably high, except for Test 2 and Test 7. Further, the four reliability estimates for Test 4 are about the same. The estimates of reliability coefficients based upon the E* group scores generally exceed those based upon the C group scores.

Table 4 presents the results of the item analysis in terms of the sign-test. For each item, the difficulty level estimate based upon the control sample scores was subtracted from the difficulty level based upon the experimental sample scores; difference greater than 1 or

---

[5]Appendix C displays item difficulty frequency distributions for each test.

Table 3

Part-test reliability (r_{tt}) estimates[a]

Test	E*	C
1	.82	.72
2	.71	.56
3	.82	.81
4-Pre[b]	.82	.80
4-Post[b]	.82	.76
5	.81	.66
6	.72	.74
7	.55	.55

[a] Computed using Kuder-Richardson formula No. 21

[b] Same test was used as Pre- and Post-test.

Table 4

Sign-test summary, complete tests

Test	Nb '+'	Nb '-'	P
1	15	7	n.s.[a]
2	23	2	< 0.01
3	17	7	n.s.
4-Pre	11	16	n.s.
4-Post	16	7	n.s.
5	18	9	n.s.
6	17	14	n.s.
7	13	12	n.s.
Total	130	74	n.s.

[a] Not significant; i.e., $p > 0.05$

Table 5

Sign-test summary: Recognition and Transfer items

Test	Multiple Choice			Transfer			Both		
	Nb '+'	Nb '-'	P	Nb '+'	Nb '-'	P	Nb '+'	Nb '-'	P
1	-	-	---	5	0	---	-	-	---
2	7	1	n.s.[a]	8	1	< 0.05	5	1	---
3	5	3	n.s.	8	1	< 0.05	1	0	---
4-Post	-	-	---	2	0	---	-	-	---
5	4	5	n.s.	2	1	---	2	1	---
6	2	3	---	10	3	n.s.	-	-	---
7	6	5	n.s.	-	-	---	-	-	---
Total	24	17	n.s.	35	6	< 0.01	8	2	n.s.

[a]Not significant; i.e., $p > 0.05$

less than -1 were counted as '+' and '-', respectively. Some of the differences were between these limits, so the number of items reported per test is not always 31.

Except for the pretest, the number of differences per test in favor of the E* group is greater. However, only on Test 2 is this statistically significant. The results for Tests 6 and 7 seem not in close agreement with those for the other tests. This may be due to the use of a random sample of students [43 from E* group, 47 from the C group] to estimate difficulty levels for each group on these two tests. On each of the other tests the entire groups were used.

Having considered the results summarized in Table 4 referring to each test in its entirely, we find it interesting to consider the data from two specific item types, recognition (multiple-choice) and transfer. Only five of the tests contained items requiring a recognition response; six of the tests contained transfer items. Table 5 gives the relevant data.

These results suggest that the treatment had little effect upon the ability to respond correctly to multiple-choice items, but had a significant overall effect upon the ability to respond correctly to transfer items, in favor of the experimental conditions. One explanation offered is that neither group had previous practice with recognition items; the Parts require constructed responses almost exclusively. On the other hand, students studying the Parts were more thoroughly and consistently exposed to "real life" situations as they learned each

new concept than is typically the case in regular classroom instruction. Thus, in a sense, they had had previous practice on transfer (type) items.

Summary

The data from the seven Part-tests provides evidence that a team of teachers, experienced in the preparation and presentation of a new mathematics curriculum [UICSM], can produce successful programed instructional materials which convey content otherwise presented only by typical classroom methods. The results reported here support the conclusion that no student was penalized in his level of mathematics achievement because of having used programed materials.

One hypothesis suggested by the data is that through the use of programed instructional materials homogeneous performance levels among students of varying abilities can be reached. This hypothesis should be checked by additional studies, at least one of which involves a longer continuous use of programed materials.

In addition, at least one such study should use programed materials designed specifically for the less able students as well as materials designed for the more able ones. It is suggested here that these materials should not be the same, branching techniques notwithstanding.

Appendix A

TOGA and STEP Results for Participating Classes in UICSM-PIP, 1961-62[a]

Raw scores

N	Group C or E	Code	School Grade	TOGA Mean	TOGA Median	STEP Mean	STEP Median
30	C	ς	8	55.77	54.00	41.90	43.00
29	E	υ	9	57.26	56.25	39.15	40.25
35	E	ι	9	53.94	55.00	40.40	42.00
26	C	ω	7	52.15	54.00	36.54	36.50
22	E	κ	7	51.96	49.00	37.00	37.50
41	E	ε	9	51.76	51.50	37.41	38.00
42	C	ο	9	48.81	49.00	35.00	35.00
46	C	φ	9	44.72	44.00	31.26	32.00
37	C	τ	9	45.14	47.00	31.14	32.50
39	E	α	9	43.79	45.00	28.89	30.00
46	E	η	9	35.60	36.00	20.96	21.00

[a] Based on Table 3, Progress Report 3 and 4, June, 1962 (see references).

Appendix B

F ratios and Associated Probability Levels, Including the Lower Ability Class

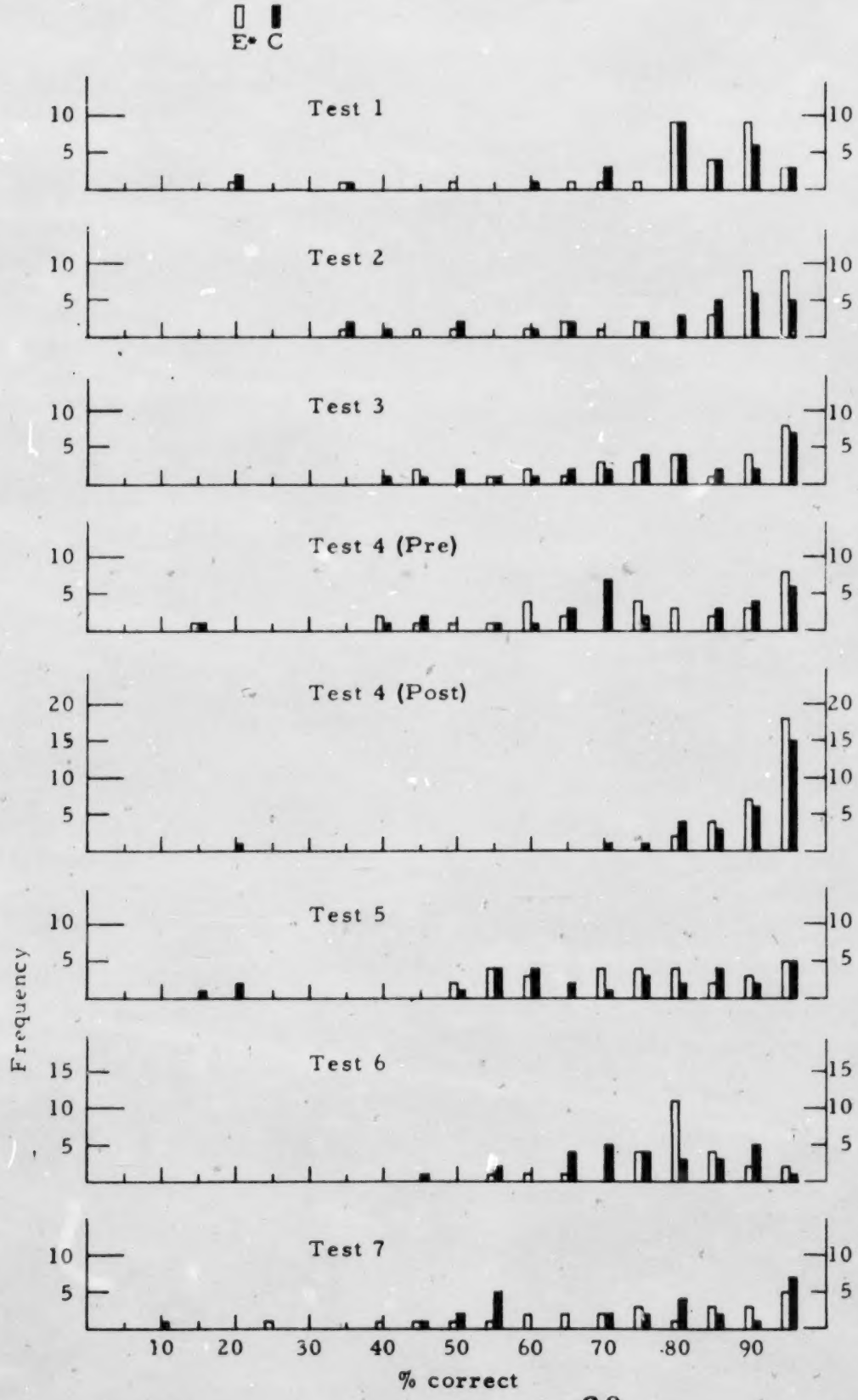
Test	Groups Compared: E and E*		E and C	
	F	p(F)	F	p(F)
General Ability				
TOGA (IQ)	1.08	---	1.52	<0.02
Math Aptitude				
STEP			2.56	<0.02
Specific Achievement				
1	1.49	<0.02	1.81	<0.02
2	1.67	<0.02	1.88	<0.02
3	1.69[a]	<0.02	1.63	<0.02
4-pre	1.57[b]	<0.02	1.77	<0.02
4-post	1.85	<0.02	2.07	<0.02
5	1.57	<0.02	2.15	<0.02
6	2.03	<0.02	1.73	<0.02
7	1.84	<0.02	1.83	<0.02

[a] Control class S failed to take this test

[b] Control class T failed to take this test

Appendix C

Item Difficulty Frequency Distributions



	Average E	Average C
Test 1	82.84	79.25
Test 2	84.26	78.13
Test 3	81.39	79.12
Test 4 (Pre)	75.08	76.11
Test 4 (Post)	94.71	90.68
Test 5	78.48	71.00
Test 6	77.92	74.40
Test 7	75.02	74.79

Appendix D

TABLE OF CONTENTS

UICSM Unit 1

Introduction

Arithmetic by mail [1-A]

Things and the names of things [1-E]

Numbers and numerals [1-K]

Part 1
43 pages

1.01 Distance and direction [1-1]

Numerals for real numbers [1-2]

Using real numbers to measure trips [1-3]

Using real numbers to locate points with respect to a given point [1-4]

Part 2
94 pages

1.02 Addition of real numbers [1-8]

Using real numbers to measure changes [1-10]

Trips of distance 0--the real number 0 [1-13]

Nonnegative and nonpositive real numbers [1-15]

Part 3
81 pages

1.03 Multiplication of real numbers [1-17]

A pump, a tank, and a movie [1-17]

Exploration Exercises--tables and operations [1-24]

1.04 Numbers of arithmetic and real numbers [1-29]

Shorter names for positive numbers [1-31]

Interpreting ambiguous numerals and words [1-31]

[CONTENTS]

1.05 Punctuating numerical expressions [1-33]

Using parentheses, brackets, and braces [1-35]

Conventions for omitting grouping symbols [1-37]

1.06 Principles for the numbers of arithmetic [1-44]

The commutative principle for multiplication [1-45]

The commutative principle for addition [1-45]

The associative principle for addition [1-46]

The associative principle for multiplication [1-46]

Another principle [1-50]

The distributive principle for multiplication over addition [1-51]

More principles [1-53]

The principle for adding 0 [1-53]

The principle for multiplying by 1 [1-53]

The principle for multiplying by 0 [1-55]

Using the principles for short cuts [1-55]

1.07 Principles for the real numbers [1-59]

Investigating the principles for the real numbers [1-60]

Using the principles for short cuts [1-62]

Part 4
78 pages

Exploration Exercises--

Subtracting undoes what adding did [1-63]

Multiplying by the reciprocal undoes what multiplying did [1-64]

Reciprocals [1-65]

[CONTENTS]

Part 5
117 pages

1.08 Inverse operations [1-66]

Operations [1-67]

Finding out what an operation is--pairs of numbers [1-67]

Finding out what the inverse of an operation is [1-68]

Subtracting a number is the inverse of adding that number [1-69]

Dividing by a nonzero number is the inverse of multiplying by that number [1-69]

Exploration Exercises--

Adding the opposite undoes what adding did [1-72]

Adding the opposite of a real number is the inverse of adding that number [1-73]

Opposites [1-75]

1.09 Subtraction of real numbers [1-75]

Subtracting a real number is the inverse of adding that number [1-75]

Subtracting is adding the opposite [1-77]

The principle for subtraction [1-79]

Changing subtraction problems to addition-of-the-opposite problems [1-79]

1.10 Opposites [1-80]

The principle of opposites [1-80]

The operation oppositing [1-81]

Using a minus sign for oppositing [1-81]

Using the principle of opposites [1-83]

[CONTENTS]

New names for negative numbers [1-86]

Three uses of the minus sign [1-87]

More names for positive numbers [1-88]

The operation "sameing" [1-88]

Three uses of the plus sign [1-88]

1.11 Division of real numbers [1-92]

Dividing by a nonzero real number is the inverse of multiplying by that number [1-92]

Ways of naming a quotient [1-94]

Numerator and denominator of a fraction [1-94]

Multiplying by 0 has no inverse [1-95]

1.12 Comparing numbers [1-95]

Comparing numbers of arithmetic [1-95]

Using the symbols '>' and '<' [1-96]

Comparing real numbers [1-97]

Part 6

78 pages

Testing by adding a positive number [1-97]

1.13 The number line [1-99]

"Lining up" the real numbers in order [1-99]

Using the symbols '$\lhd$' and '$\rhd$' [1-100]

Using the symbols '$\geq$' and '$\leq$' [1-101]

Uniform scale [1-102]

Absolute value operation

Distance between real numbers [1-103]

Part 7

91 pages

Absolute value of a real number [1-104]

Using absolute value bars [1-106]

Does absolute valuing have an inverse? [1-107]

Operations and their inverses [1-108]

Ambiguous numerals [1-110]

References

1. Beane, Donald. A comparison of linear and branching techniques of programed instruction in plane geometry, U. S. Office of Education, Title VII, Project 711151-01, Comparative studies of principles for programing mathematics in automated instruction, Technical Report No. 1, Urbana: University of Illinois, July 1962.
2. Smith, Leone. Programed learning in elementary school: an experimental study of relationships between mental abilities and performance, U. S. Office of Education, Title VII, Project 711151-01, Comparative studies of principles for programing mathematics in automated instruction, Technical Report No. 2, Urbana: University of Illinois, August, 1962.
3. Semi-Annual Reports No.'s 3 and 4, Comparative studies of principles for programing mathematics in automated instruction, Period: December 6, 1961 - June 6, 1962.
4. UICSM High School Mathematics, Unit 1, Urbana; University of Illinois Press, 1960.